Signal Extraction in Camera-Based Photoplethysmography Using a Modified Wiener Filter

Alexander Trumpp, Hagen Malberg und Sebastian Zaunseder

Institute for Biomedical Engineering, TU Dresden, Dresden, Germany
Contact: alexander.trumpp@tu-dresden.de

Introduction

Camera-Based Photoplethysmography (cPPG) is a remote measuring technique which allows the acquisition of cardiovascular signals using cameras [1]. The blood volume pulse causes pulsatile skin color changes that are not visible but detectable by camera sensors. Video recordings allow the spatio-temporal capture of these changes. It was shown that cPPG can be applied in the human face [2, 3, 4], using regular video cameras or webcams, respectively. Such characteristics render the technique a promising mean for vital parameter acquisition in daily life situations.

For signal (plethysmogram) extraction, a region of interest (ROI), typically a rectangular area, is defined and its pixel values are averaged (Arithmetic Mean Filter – AM) per frame. Due to appearing image noise and physiological factors, however, the pixels' information content varies. Against that background, averaging must be questioned as optimal mean of processing.

Within this contribution, we propose a modified Wiener filter (WM) that considers the ROI content and adapts to certain variations of pixel values. We test its ability to improve the PPG signal quality compared to AM.

Methods and Material

Test Subjects and Instrumentation

25 healthy subjects (24.3 ± 3.5 years) were recorded in a lying position using an IDS UI-5240CP-C-HQ RGB camera at a resolution of 320x240 pixels, a color depth of 3x10 Bit and a frame rate of 100 fps. The videos obtained subjects' frontal face during a vigilance task that led to slight movements. The experiments were executed indoors where a fluorescent light source was used as illumination. For further analyses, we only processed videos green channel since it was shown to provide the best signal qualities [3].

As reference, we placed a PPG Sensor (ADInstruments IR Plethysmograph MLT1020EC) at each subject's right ear and acquired a plethysmogram using ADIntruments' PowerLab 16/35.

Image Processing

From each subject, ten subsequent intervals of 15 s were considered. At the beginning of each interval, the face was detected by an algorithm which is based on the work of Viola and Jones [5]. It returns a box around the face, however, only the center of 60 % box' width was used as ROI. Beside the face ROI, two square ROI on the cheeks and one on the forehead were set automatically, relative to the size of the detected box. It was shown that ROI at these locations can provide higher signal qualities than a larger face ROI [6]. For each 15 s interval, the ROI location kept unchanged. The resulting ROI sizes were around 138x82 pixels for the face, 27x27 pixels for the forehead and 24x24 pixels for the cheeks. Figure 1 shows an example for the ROI positioning.

The Wiener theory [7] builds the foundation for signal and image filters. Assuming that an image $f(x,y)$ is degraded by white additive zero-mean and signal-independent noise, the Wiener filter for image noise reduction can be written as [8, eq. 9.23 -modified]:

$$w(u,v) = \frac{\sigma_f(u,v)^2}{\sigma_f(u,v)^2 + \sigma_n(u,v)^2},$$

where (u,v) denote an image sub-region, σ_f the variance in $f(u,v)$ calculated around a neighborhood NxN for every pixel position, and σ_n the noise variance. We determined σ_n as mean variance of σ_f which, itself, is estimated with the variance σ_g of the degraded (measured) image region $g(u,v)$ [8, eq. 9.29a -modified]:

$$\sigma_f = \begin{cases} \sigma_g(u,v)^2 - \sigma_n{}^2, & if\ \sigma_g(u,v)^2 > \sigma_n{}^2 \\ 0, & otherwise \end{cases}.$$

Using the Wiener filter formulation, we built a weighting matrix for the ROI (u,v) of every frame which reads to:

$$w_{ROI}(u,v) = 1 - w(u,v)\ .$$

Figure 1 shows an illustration of w_{ROI} for the face, the forehead and both cheeks of one frame.

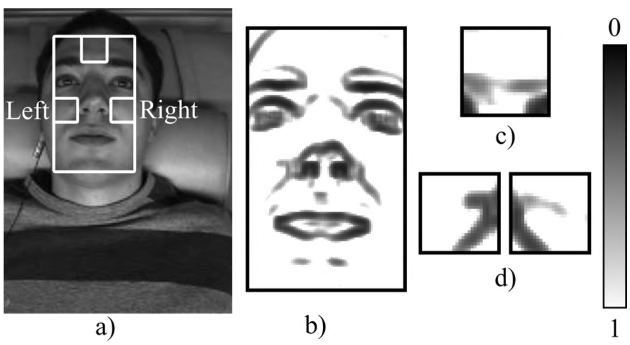

Figure 1: Example for automatic placed ROI (a) at face (outer border), forehead and both cheeks, and their calculated weighting matrixes (b-face, c-forehead, d-cheeks)

Signal Extraction and Filter Parameterization

To finally achieve the cPPG signal, a weighted mean filter was applied in every frame ROI using w_{ROI}. We tested six different neighborhood sizes N (filter orders), three adaptive sizes: $(^1/_5 \text{ x } ^1/_5)b$, $(^1/_4 \text{ x } ^1/_4)b$, $(^1/_3 \text{ x } ^1/_3)b$, where b is the ROI width (WM1, WM2, WM3), and three static sizes: 3x3, 5x5, 7x7 (WM4, WM5, WM6). Due to the different ROI locations and filter orders 240 cPPG signals of 15 s were extracted for every subject. For comparison, AM was applied as well (40 signals per subject).

Quality Assessment

Our method was evaluated in terms of heart frequencies' detection rate (HDR) and signal-to-noise ratio (SNR). Both parameters were calculated by means of the amplitude spectrum for each 15 s detrended and windowed signal. The heart rate (HR) was detected at the maximum amplitude between 0.5 and 4 Hz. A HR is considered as correct (HDR=100 %) when it doesn't deviate more than ±5 bpm from the reference HR which is calculated out of the reference plethysmogram.

The SNR was derived based on a formulation of de Haan [9, eq. 19], considering a spectral band of ±5 bpm around the detected heart frequency and its first harmonic as signal component and the residual amplitudes within [0.5, 4] Hz as noise.

Results

Figure 2 shows the results for HDR after averaging all interval and subject HDR values for each method and configuration, respectively. The highest HDR were achieved in the face ROI, where WM outperformed AM (95.6 % vs. 93.2 %). For the other ROI locations, WM and AM yielded similar rates (forehead: 92.0 % vs. 92.4 %, left cheek: 90.8 % vs. 91.2 %, right cheek: 92.0 % vs. 91.6 % - considering the best filter order setting).

Using a 2-way repeated measures ANOVA (factors: ROI and method), both factors proved to significantly affect the SNR. Calculating SNR's estimated marginal means (Figure 3), WM performed best in the face ROI (4.05 dB vs. AM: 2.74 dB). Results for the other ROI locations were similar (WM vs. AM - forehead: 3.88 dB vs. 3.89 dB, left cheek: 3.95 dB vs. 3.89 dB, right cheek: 3.79 dB vs. 3.81 dB).

The choice of the filter order had minor effect for the forehead and both cheeks. For the face ROI, the configuration using an adaptive filter order at factor $^1/_5$ (WM1) seem to significantly enhance the outcome. Applying static orders, larger sizes achieve better results.

Discussion

The estimated noise variance is recalculated for every ROI and depends strongly on the ROI content. High local variances lead to a higher value σ_n, and hence, to higher weightings for pixels in regions with lower variances. Therefore, e.g. in the face ROI, regions around the eyes, the nose or the mouth are less considered for the signal extraction which result in an increase of the HDR and SNR. This adaption is an advantage for bigger and more heterogeneous ROI.

For homogeneous ROI - ideally forehead and cheeks - the noise variance is smaller and more pixels might be weighted lower. In that case, image noise should affect w_{ROI} stronger (lower weightings). Nevertheless, clear improvements for WM were not observed there.

The choice of the filter order plays a key role for the face ROI. A proper setting for WM can even achieve better signal qualities (SNR) as the forehead and the cheeks yield, which were reported as most capable regions by Lempe et al. [6].

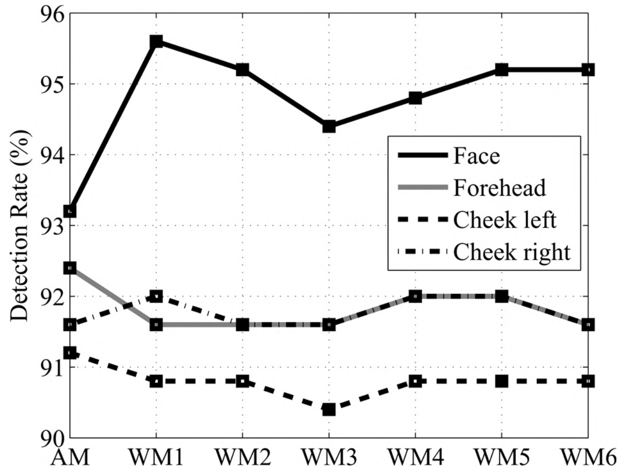

Figure 2: Results for heart frequencies' detection rate for different ROI (face, forehead, both cheeks) using the arithmetic mean filter (AM) and the modified Wiener filter (WM) with six different order configurations

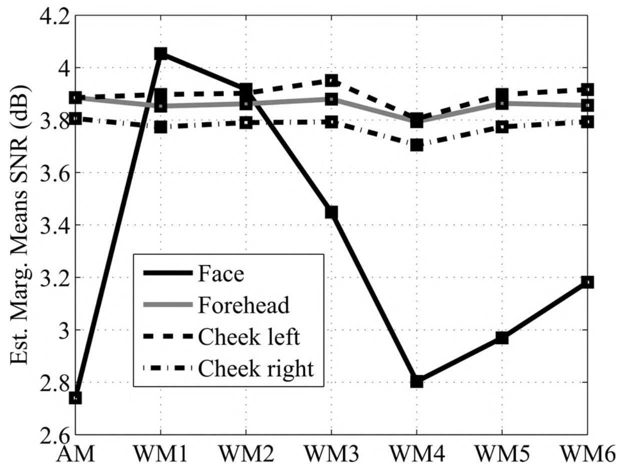

Figure 3: Results for SNR's estimated marginal means for different ROI (face, forehead, both cheeks) using the arithmetic mean filter (AM) and the modified Wiener filter (WM) with six different order configurations

Conclusion

Our results prove that the modified Wiener filter is an appropriate mean to improve signal quality as well as heart rates detection rate and is a preferable alternative to standard averaging. The determination of the filter order, as well as the ROI noise variance has to be evaluated in the future. A static value for σ_n could be a better choice than a ROI adaption.

References

[1] HUELSBUSCH, M. ; BLAŽEK, V.: Contactless mapping of rhythmical phenomena in tissue perfusion using PPGI. In: *Proc. SPIE 4683, Medical Imaging 2002: Physiology and Function from Multidimensional Images.* San Diego, February 2002, pp. 110-117

[2] TAKANO, C. ; OHTA, Y.: Heart rate measurement based on a time-lapse image. In: *Medical engineering & physics 29* (2007), Nr. 8, pp. 853–857

[3] VERKRUYSSE, W. ; SVAASAND, L. ; NELSON, J.: Remote plethysmographic imaging using ambient light. In: *Optics Express 16* (2008), Nr. 26, pp. 21434-21445

[4] POH, M. ; MCDUFF, D. ; PICARD, R.: Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. In: *Optics Express 18* (2010), Nr. 10, pp. 795–805

[5] VIOLA, P. ; JONES, M.: Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.* Kauai, December 2001, pp. I-511–I-518

[6] LEMPE, G. ; ZAUNSEDER, S. ; WIRTHGEN, T.: ROI Selection for Remote Photoplethysmography. In: *Bildverarbeitung für die Medizin 2013,* Springer Berlin Heidelberg, 2013, pp. 99-103

[7] WIENER, N.: *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. MIT Press Cambridge, 1949

[8] LIM, J.: *Two-Dimensional Signal and Image Processing*. Prentice Hall, 1990, pp. 527-540

[9] DE HAAN G., ; JEANNE V.: Robust pulse-rate from chrominance-based rPPG. In: *Biomedical Engineering, IEEE Transactions on 60* (2013), Nr. 10, pp. 2878-2886